

【关联工单】

【L5】【8701278】EMR 通过控制台修改impala 的mem_limit参数，修改无法成功

【复现步骤】

/data/Impala/bin/impala-shell/impala-shell -i 10.3.59.43:27001

数据库: oneid

复现sql

```
use oneid;
select t2.t_id,
       coalesce(t1.subscriber_id,t2.subscriber_id) as subscriber_id,
       coalesce(t1.user_id,t2.user_id) as user_id,
       t1.device_id as device_id,
       t1.device_id_type as device_id_type,
       t1.openid as openid,
       t1.union_id as union_id,
       coalesce(t1.phone,t2.phone) as phone,
       coalesce(t1.phone_md5,t2.phone_md5) as phone_md5,
       coalesce(t1.user_first_version,t2.user_first_version) as user_first_version,
       coalesce(t1.user_first_product,t2.user_first_product) as user_first_product,
       t1.sequece_id,
       t1.source_platform as source_platform,
       t1.insert_date as insert_date,
       t1.update_date as update_date,
       t1.batch_date as batch_date,
       coalesce(t1.user_type,t2.user_type) as user_type,
       t1.get_through_label as get_through_label
from merge_day_tid_device_data_incr t1
left join (select a.t_id,a.subscriber_id,a.user_id,a.phone,a.phone_md5,a.user_first_version,a.user_first_product,a.user_type,a.device_id
           from tid_device_user_data a
           join test on 1=1
           where a.batch_date not in ('20220425')) t2 on t1.device_id = t2.device_id;
```

【背景】

客户2022年4月26号反馈 如上sql，执行过程中必须加上set mem_limit=1g;后能成功，不加就会失败，profile信息如附件:[profile.txt](#)

【排查】

2022年4月26日，排查了impalad日志，未发现异常，查看机器资源未发现异常，连接数未发现异常

2022年4月27日，获取到客户测试环境信息，多次进行复现操作，排查日志，并未在日志中发现报错的具体原因

2022年4月28日，再次登录客户环境复现排查，同时观察机器内存使用情况，发现机器内存非常少只有1-2g左右了，于是调整了全局的mem_limit降低到50%，在次测试，问题还是复现，查看sql执行的profile信息，有如下输出：

```
EXCHANGE_NODE (id=7) could not allocate 512.00 KB without exceeding limit.
Error occurred on backend 10.3.59.43:27002
Memory left in process limit: 2.77 GB
```

【原因总结】

EXCHANGE_NODE 使用的内存不在mem_limit=80%控制范围内，而是机器上剩余的内存，当mem_limit的值越大，留给exchange的内存就会越小，出现这个的报错的概率就会变大

impala sql运行中内存使用分为 tracker memory pool 和untrack memory pool.tracker memory pool 对应的是全局的mem_limit, untracker memory pool是机器剩余可用物理内存, 当mem_limit设置较大时, untracker memory pool 就会变小, untracker memory pool主要是用作 shuffle, udf等使用, 不在mem_limit控制范围内。

当mem_limit在session级别设的太小会导致如下错误,也就是tracker memory pool设置过小导致的, 错误信息如下:

```
Query submitted at: 2022-04-28 19:31:09 (Coordinator: http://10.3.59.43:27004)
Query progress can be monitored at: http://10.3.59.43:27004/query_plan?query_id=014752b019c489fc:25cf5e6a00000000
ERROR: Memory limit exceeded: Error occurred on backend 10.3.59.43:27002 by fragment 014752b019c489fc:25cf5e6a000000009
Memory left in process limit: 21.08 GB
Memory left in query limit: -156.09 KB
Query(014752b019c489fc:25cf5e6a00000000): memory limit exceeded. Limit=100.00 MB Reservation=34.00 MB ReservationLimit=68.00 MB OtherMemory=66.15 MB Total=100.15 MB Peak=100.15 MB
  Fragment 014752b019c489fc:25cf5e6a00000000: Reservation=0 OtherMemory=25.75 KB Total=25.75 KB Peak=413.00 KB
```

【建议】

遇到如上错误时需要手动指定mem_limit参数, 具体值需要根据业务确定
如: set mem_limit=10g;